MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

AD-A249 799

PAGE

Form Approved OMB No. 0704-0188

Public reporting ... per response, including the time for reviewing instructions, searching existing data sources, gathering and ... of information. Send comments regarding this burden estimate or any other aspect of this collection of inf... Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, S... and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY	March 1992	3. REPORT TYPE AND DATES COVERED Final 15 Jul 88-31 Dec 91

4. TITLE AND SUBTITLE Topics in the Sequential Design of Experiments	5. FUNDING NUMBERS DAAL03-88-K-0122
6. AUTHOR(S) Michael Woodroofe	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Michigan Ann Arbor, Michigan 48109	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING/MONITORING AGENCY REPORT NUMBER ARO 25819.9-MA

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

This report contains abstracts of publications and a listing of dissertations.

14. SUBJECT TERMS Design of Experiments, Renewal Theory, Sequential Testing Limit Theory, Local Limit Theorem	15. NUMBER OF PAGES 12 16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

ARO 25819.9-MA

FINAL REPORT ON

TOPICS IN THE SEQUENTIAL DESIGN OF EXPERIMENTS

DAAL 03 88 0122

Author of Report: Michael Woodroofe (Principle Investigator)
Statistics Department
University of Michigan
Ann Arbor, Michigan 48109

Publications

The following papers were written with the support of the grant and have appeared or are scheduled to appear. Abstracts of the papers are appended

[1] "On the non linear renewal theorem," by Michael Woodroofe. *Ann. Prob.*, 18 (1990), 1790-1805.

[2] "Stopping times and stochastic monotonicity," by Michael Woodroofe. *Sequential Analysis*, 9 (1990), 335-342.

[3] "Adaptive biased coin designs for the Behrens Fisher Problem," by Jeffrey Eisele. *Sequential Analysis*, 9 (1990), 343-360.

[4] "One armed bandit problems with covariates," by Jyoti Sarkar. *Ann. Statist.*, 19 (1991), 1078-2002.

[5] "Integrable expansions for posterior distributions for one parameter exponential families," by Michael Woodroofe. *Statistica Sinica*, 2 (1991), 91-112.

[6] "A non linear renewal theory for a functional of a Markov random walk," by Vince Melfi. To appear in *Ann. Prob*, April 1992.

[7] "Estimation following sequential testing: a simple approach for a truncated S.P.R.T.. To appear in *Biometrika*.

Submitted Papers

The following papers have been submitted for publication. Abstracts of these papers are appended.

[8] "A local limit theorem for perturbed random walks," by Mei Wang. Submitted to *Statistics and Probability Letters*.

[9] "Non parametric Bayes estimation of a distribution function with truncated data," by Mauro Gasparini. Submitted to *J. Statist. Plan. Inf.*.

[10] "The doubly adaptive biased coin design," by Jeffrey Eisele. Submitted to *J. Statist. Plan. Inf.*

Technical Report

"A martingale proof of normality for Wei's biased coin design," by Michael Woodroofe and Jeffrey Eisele, 1990. Statistics Department, University

Dissertations

The following students completed dissertations with the support of the grant.

Jyoti Sarkar (1990). *Bandit Problems with Covariates: Sequential Allocation of Experiments*. Statistics Department, University of Michigan.

Mei Wang (1990). *Local Limit Theorem and Occupation Times for Perturbed Random Walks*. Mathematics Department, University of Michigan.

Vince Melfi (1991). *Nonlinear Markov Renewal Theory with Applications to Sequential Analysis*. Statistics Department, University of Michigan.

Jeffrey Eisele (1991). *Doubly Adaptive Biased Coin Designs*. Statistics Department, University of Michigan.

Mauro Gasparini (1992, expected). *Topics in Bayesian non parametric inference* (tentative). Statistics Deparment, University of Michigan.

Research Findings

The main research findings are described below in the abstracts of the papers cited above.

The Annals of Probability
1990, Vol. 18, No. 4, 1790–1806

ON THE NONLINEAR RENEWAL THEOREM[1]

BY MICHAEL WOODROOFE

The University of Michigan

Let $Z_1, Z_2, \dots$ be jointly distributed random variables for which $\sup_k Z_k = \infty$ w.p.1 and let $t = t_a = \inf(n \ge 1: Z_n > a)$ and $R_a = Z_t - a$ for $a \ge 0$. Conditions under which R_a has a limiting distribution as $a \to \infty$ are developed. These require that the finite dimensional, conditional distributions of the increments $Z_{t+k} - Z_t$, $k \ge 1$, converge to the finite dimensional distributions of a process for which the result is known, thus weakening the slow change condition in earlier work. The main result is applied to some sequences for which the limiting distributions are those of the partial sums of an exchangeable process. These include the Euclidean norms of a driftless random walk in several dimensions and sequences for which the conditional distribution of $Z_{n+1} - Z_n$ given the past has a limit w.p.1 as $n \to \infty$.

1. Introduction. Let $(\Omega, \mathscr{A}, P)$ denote a probability space, let $\mathscr{A}_1 \subseteq \mathscr{A}_2 \subseteq \dots$ denote subsigma-algebras of $\mathscr{A}$ and let $Z_1, Z_2, \dots$ be random variables, defined on $(\Omega, \mathscr{A}, P)$, for which Z_k is $\mathscr{A}_k$ measurable for all $k = 1, 2, \dots$ and

$$\sup_{k \ge 1} Z_k = \infty \quad \text{w.p.1.} \tag{1}$$

Such a sequence may be called an *infinite supremum process*. For any such process, the first passage times and excesses

$$t_a = \inf\{k \ge 1: Z_k > a\} \tag{2}$$

and

$$R_a = Z_{t_a} - a, \tag{3}$$

may be defined for all $a \ge 0$ w.p.1. Let H_a denote the distribution function of R_a; that is,

$$H_a(r) = P\{t_a < \infty, R_a \le r\}, \qquad \forall\, a, r \ge 0. \tag{4}$$

The problem considered is to find conditions under which R_a has a limiting distribution H as $a \to \infty$; that is, $H_a \Rightarrow H$ as $a \to \infty$, where $\Rightarrow$ denotes weak

Received May 1988; revised June 1989.
[1]Research supported by the National Science Foundation under Grant DMS-84-13452 and the U.S. Army Research Office under Grant DAAL03-88-0122.
AMS 1980 *subject classification*. 60K05.
Key words and phrases. First passage times, excess over the boundary, limiting distribution, random walks, exchangeable processes.

The Annals of Statistics
1991, Vol. 19, No. 4, 1978–2002

ONE-ARMED BANDIT PROBLEMS WITH COVARIATES[1]

BY JYOTIRMOY SARKAR

University of Michigan

As does Woodroofe, we consider a Bayesian sequential allocation between two treatments that incorporates a covariate. The goal is to maximize the total discounted expected reward from an infinite population of patients. Although our model is more general than Woodroofe's, we are able to duplicate his main result: The myopic rule is asymptotically optimal.

1. Introduction.

1.1. *Statement of the problem.* Consider a population of patients who arrive sequentially for treatment of a disease. Suppose each patient may be treated with either a standard treatment whose statistical characteristics are known, or a new treatment whose characteristics are unknown. Also suppose that before deciding to assign a given patient to a treatment, we observe a covariate X, such as age, severity of disease or general physical status, which is specific to the patient.

Let Y^0 and Y^1 denote the potential rewards from the standard and the new treatment, respectively. Let $\delta = 0$ and $\delta = 1$ denote the choice of standard or new treatment, respectively. We would like to assign patients to treatments in such a manner that the total discounted expected reward over the whole population of patients is maximized. The discount sequence is geometric with discount factor $\alpha \in (0, 1)$.

1.2. *Rationale for covariate model.* In clinical trials, the goals an experimenter would like to attain are diverse and often conflicting. Ethical considerations are prominent in all experimentation involving human subjects. Conflicts are invariably generated by the obligation of a researcher to balance the well-being of the current patient (individualistic view) with that of the future patients (utilitarian view) who stand to benefit from new advances in medical treatment. This long-standing dilemma has received considerable attention in both statistical and medical literature such as Anscombe (1963), Weinstein (1974), Byar, Simon, Friedewald, DeMets, Ellenberg, Gail and Ware (1976), Bartlett, Roloff, Cornell, Andrews, Dillon and Zwischenberger (1985) and Woodroofe and Hardwick (1990).

Received November 1989; revised January 1991.

[1]Research supported by U.S. Army Research Office under DAAL 03-88-0122 and NSF under DMS 89-02188.

AMS 1980 *subject classification*. 62L10.

Key words and phrases. Sequential allocation, one-armed bandit problem, Bayesian analysis, regret, myopic policy, asymptotically optimal.

SEQUENTIAL ANALYSIS, 9(4), 335-342 (1990)

ON STOPPING TIMES AND STOCHASTIC MONOTONICITY

By

Michael Woodroofe

The University of Michigan

Key words and phrases: *estimation after sequential testing; exponential families; the fundamental identity of sequential analysis.*

ABSTRACT

If t is any stopping time, then the distributions of $\bar{X}_t$ are stochastically monotone in the parameter, when sampling from a one parameter exponential family. The proof is a simple exercise in differentiating the fundamental identity of sequential analysis. Some other applications of the technique are included.

1. INTRODUCTION..

Siegmund (1978) has suggested an ingenious method for setting confidence intervals after sequential testing, which involves ordering boundary points in a counter clockwise direction and then inverting a collection of tests. See also Siegmund (1985, Sections 3.4 and 4.5). Recently, Bather (1988) has derived a stochastic monotonicity property for the counter clockwise ordering for generalized sequential probability ratio tests and exponential families.

An alternative approach to estimation after sequential testing is to estimate the parameter by its maximum likelihood

INTEGRABLE EXPANSIONS FOR POSTERIOR DISTRIBUTIONS FOR ONE-PARAMETER EXPONENTIAL FAMILIES

Michael Woodroofe

University of Michigan

Abstract: **The main results provide asymptotic expansions for posterior distributions which may be integrated termwise with respect to the marginal distribution of the data. The proof uses a data dependent transformation which converts the likelihood function to exact normality and then applies a version of Stein's Identity to the posterior distributions. Applications to sequential confidence intervals are described briefly.**

Key words and phrases: **Posterior distributions, parameter transformations, Stein's Identity, martingale convergence theorem, stopping times, sequential confidence intervals.**

1. Introduction

Asymptotic expansions for posterior distributions may be traced to the time of Laplace, but only recently have researchers investigated conditions under which these expansions may be integrated with respect to the marginal distribution of the data. See Johnson (1970) for a rigorous account of the pointwise expansions and Bickel, Goetze and Van Zwet (1985) and Ghosh, Sinha and Joshi (1983) for recent work on integrating them.

In this article, an alternative approach to the expansions is presented which makes the question of integrability more transparent. First, instead of renormalized estimation error, a data dependent transformation (3) of the parameter is considered, which converts the likelihood to exact normality. Then a version of Stein's Identity is applied to the posterior distributions to isolate the remainder terms. The alternative approach avoids the use of messy Taylor series expansions and leaves the renormalized remainder terms in the form of conditional expectations, so that the martingale theory may be brought to bear on the integrability question.

Integrable expansions for posterior distributions are needed in design problems where the overall Bayes risk must be computed in order to see the effect

AN ADAPTIVE BIASED COIN DESIGN FOR THE BEHRENS-FISHER PROBLEM

Jeffrey R. Eisele

Department of Statistics
The University of Michigan
Ann Arbor, MI 48109

Key words and Phrases: asymptotic optimality; experimenter bias; sampling plan; sequential design; stopping rule.

ABSTRACT

Wei(1978) introduced the adaptive biased coin design to reduce experimenter bias and offer a compromise between perfect balance and complete randomization. In situations such as the Behrens-Fisher problem, balance is not necessarily desired and the optimal ratio of sample sizes is unknown. To reduce experimenter bias, by introducing randomization, an adaptive biased coin design is superimposed on Robbins, Simons, and Starr's(1967) sequential analogue of the Behrens-Fisher problem. The design has asymptotic properties similar to Robbins, Simons, and Starr's sequential procedure.

1. INTRODUCTION

Suppose patients arrive sequentially at an experimental site and are assigned immediately to one of two treatment groups A or B. It is desired to design a sequential procedure, with a randomized allocation scheme, for the fixed width interval estimation of the difference of the means of these two populations.

NONLINEAR MARKOV RENEWAL THEORY WITH STATISTICAL APPLICATIONS

By

Vincent F. Melfi

The University of Michigan

ABSTRACT

An analogue of the Lai-Siegmund nonlinear renewal theorem is proved for processes of the form $S_n + \xi_n$, where $\{S_n\}$ is a Markov random walk. Specifically, $Y_0, Y_1, \ldots$ is a Markov chain with complete separable metric state space; $X_1, X_2, \ldots$ is a sequence of random variables such that the distribution of X_i given $Y_j,\ j \geq 0$ and $X_j,\ j \neq i$ depends only on Y_{i-1} and Y_i; $S_n = X_1 + \cdots + X_n$; and $\{\xi_n\}$ is slowly changing, in a sense to be made precise below. Applications to sequential analysis are given with both countable and uncountable state space.

Key words and phrases. Markov chain, Markov random walk, nonlinear renewal theorem, excess over the boundary, repeated significance test.

AMS(1980) Classifications. Primary 60K05, 60K15; secondary 62L10, 60J10.

Research supported by the U.S. army under DAAL 03-88-0122

Imp Col Biometrika Batch 671 Galley 1 K31
Biometrika (1992), 79, 0, pp. 000–00
Printed in Great Britain

Estimation after sequential testing: a simple approach for a truncated sequential probability ratio test

By MICHAEL WOODROOFE
Statistics Department, The University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

Summary

An approximate pivot is constructed for the problem of estimating a normal mean θ following a truncated sequential probability ratio test and shown to provide a useful method for constructing confidence bounds and intervals. Letting t denote the sample size and S_t the sum of observations, the approximate pivot is constructed by standardizing $S_t^* = t^{-\frac{1}{2}}(S_t - t\theta)$ the mean and variance of which are no longer 0 and 1, due to the optional stopping. The truncation of the sequential probability ratio test is done in a nonstandard way in order to smooth the boundary.

Some key words: Asymptotic normality; Confidence levels; Error probabilities; Expected sample size; Means and variances; Simulation.

1. Introduction

Sequential tests have long been advocated as a means of reducing the ethical problems inherent in randomized clinical trials on human patients. There are now several good sequential tests which are finding increasingly widespread applications. See, for example, the books by Armitage (1975), Siegmund (1985) and Whitehead (1983) for complementary accounts of these developments.

Of course, the use of a sequential test complicates the problem of estimating parameters, after the test has been concluded. In response to this problem, Armitage (1957) proposed a confidence procedure which explicitly incorporates the stopping rule into the calculation of coverage probabilities for binary data; and the proposal was developed in some detail by Siegmund (1978) for normal data. In effect, Armitage and Siegmund order the points on the stopping boundary in a natural way; they then use this ordering to construct a family of tests of hypotheses of the form $\theta \leq \theta_0$ for arbitrary θ_0; and then they invert this family of tests to form confidence bounds. See also Siegmund (1985, §§ 3.4, 4.5). This approach has attracted substantial interest recently (Bather, 1988; Facey & Whitehead, 1990; Kim, 1987). On the other hand, it is moderately complicated.

The purpose of the present paper is to present an alternative confidence procedure, which starts with an approximately pivotal quantity and then proceeds in natural ways. The alternative seems both conceptually and technically simpler.

The alternative is developed in the context of a nonstandard truncated sequential probabability ratio test about a normal mean, although the method is fairly general. Truncated sequential probability ratio tests are described in § 2, and the alternative approach in § 3. In § 4 the results of a simulation study of the alternative procedure are reported. Section 5 contains some brief remarks on possible modifications of the alternative procedure and its domain of applicability; § 6 contains a rough derivation of an approximation to the error probabilities.

Nonparametric Bayes Estimation of a distribution function with truncated data

Mauro Gasparini[1]

March 20, 1991

The University of Michigan

Abstract

A truncation bias affects the observation of a pair of variables (X,Y), so that data are available only if $Y \leq X$. In such a situation, the nonparametric maximum likelihood estimator (NPMLE) of the distribution function of Y may have unpleasant features (Woodroofe (1985)). As a possible alternative, a nonparametric Bayes estimator is obtained using a Dirichlet-Ferguson prior. Its frequentist asymptotic behavior is investigated and found to be the same as the asymptotic behavior of the NPMLE. The results are illustrated by an example, with astronomical data, where the NPMLE is clearly unacceptable.

[1]Research partially supported by U.S.Army Research Office under DAAL-03-88-0122.

[0]*Key words and phrases.*Nonparametric, Bayes estimation, Dirichlet-Ferguson prior, truncated data, luminosity, quasar.

[0]*AMS 1980 subject classifications.*Primary 62G05;secondary 62P99.

THE DOUBLY ADAPTIVE BIASED COIN DESIGN

Jeffrey R. Eisele*

Biostatistics Center, University of Wisconsin, Madison, WI 53706

Abstract: A biased coin design is proposed for the allocation of subjects to treatments A and B in a clinical trial when the desired allocation proportions are unknown. The design is doubly adaptive in the sense that it takes account of the proportion of subjects assigned to each treatment and the current estimate of the desired allocation proportion. A strong law of large numbers is established for the proportion of subjects assigned to a treatment when subject responses are independent random variables from standard exponential families. The normal case is presented as an application.

AMS Subject Classification: Primary 62L05; Secondary 60F15.

Key words and phrases: Biased coin designs, exponential families, martingale, sequential procedure, stopping rule, strong law of large numbers.

* Research supported by the U.S. Army under DAAL-03-88-0122.

A LOCAL LIMIT THEOREM FOR PERTURBED RANDOM WALKS

By Mei WANG*

Department of Mathematics, The University of Michigan, Ann Arbor, MI

Abstract: The main result reported here is a Stone type local limit theorem for perturbed random walks $Z_n = S_n + \xi_n$ when some slow variation conditions are imposed on ξ_n's.

Key words: Local (central) limit theorem, Edgeworth expansion, perturbed random walks.

1. Background

When specialized to one dimension and non-lattice distributions, Stone's Theorem (1965) asserts the following: Let $X_1, X_2, X_3, \cdots$ denote i.i.d. non-lattice random variables for which $\mathbb{E}[X_i] = 0$ and $\mathbb{E}[X_i^2] = 1$. Let

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad n = 1, 2, 3, \cdots$$

be the sums, called "random walks". Then for each $L \in [0, \infty)$,

$$\varepsilon_n(L) = \sup_{c \le L} \sup_{b \in \mathbb{R}} \left| \sqrt{n}\,\mathbb{P}\{b < S_n \le b + c\} - c\,\phi\left(\frac{b}{\sqrt{n}}\right) \right| \underset{n\to\infty}{\longrightarrow} 0,$$

where $\phi(x)$ is the standard normal density and $\mathbb{R}$ is the real line. A consequence of this is

$$\sqrt{n}\,\mathbb{P}\{S_n \in J\} \underset{n\to\infty}{\longrightarrow} \frac{|J|}{\sqrt{2\pi}}$$

for any interval $J \subset \mathbb{R}$ of length $|J|$. This is the result of Shepp (1964).

There has been recent interest in sequences of random variables, called perturbed random walks. See, for example, Siegmund (1985), Woodroofe (1982), and Lalley (1984). A

* Research partially supported by the U.S. Army Research Office under DAAL 03-88-K-0122.